WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

HIERARCHICAL PRODUCTION PLANNING

Harlan C. Meal

WP # 1497-83 11/1/83

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

HIERARCHICAL PRODUCTION PLANNING

Harlan C. Meal

WP # 1497-83 11/1/83

An information system that is inappropriate for the manufacturing operation where it is installed can be a source of serious difficulties. In one case, the mismatch between the needs of the operation and the characteristics of the manufacturing informatiom system nearly pushed a company into bankruptcy.

The electronics instruments produced by the company were changing rapidly. New versions of older instruments were frequently introduced and new ways of performing functions would lead to entirely new products.

A single instrument could be built in many different variations. There were several different input and output connectors and several voltage and frequency possibilities for power. In addition, the panels could be customized to meet a particular customer's needs or desires.

In the face of the rapid rate of product change and a company practice of tailoring products to meet the customer's desires, the firm attempted to install and operate a material requirements planning system with a planning horizon equal to the longest of its parts procurement and assembly lead times - 56 weeks (16 weeks for manufacturing, 40 weeks for the longest parts procurement).

The system was regenerated (updated) every four weeks. The sales department, under protest, provided fifty-six weeks of sales projections, by item, including custom variations. Having done that, they insisted on being allowed to change their projections when the system was updated four weeks later. At that updating, it was apparent that part of the new master schedule was not feasible, that the needed parts could not be obtained in time to allow normal manufacturing lead times. At the insistence of the sales department, the manufacturing people agreed to do their best to expedite the parts and complete the revised master schedule on time.

At the second updating, eight weeks after the start, the scenario was repeated. There were more changes in the master schedule, more

infeasibilities, and more friction between sales and manufacturing. This situation continued at further updatings of the system.

As time went on, part shortages became endemic. Investments increased in inventories tied up because needed parts were not available. Worker efficiency went down because of inability to complete an assembly task. Part deliveries were delayed because the company ran short of capital (because of excess inventory investment) and stretched out its payment schedules. Suppliers responded with credit holds on parts orders.

It became clear that the long planning horizon required by the MRP system was completely inappropriate for a company of this sort which had a volatile product line and a partially customized product. Their own marketing policies would simply not allow them to hold a master schedule frozen for more than a few weeks.

How could they have made such a mistake? Why did they install this system? The answer was easy to find. They installed the new system because it promised to reduce rapidly growing inventories of parts and subassemblies The new system not only failed to halt the inventory expansion, but nearly resulted in bankruptcy. An inappropriate manufacturing information system can have serious, even fatal, consequences.

This firm in only an example of the many firms in which large and expensive production planning and control system projects have been less than resoundingly successful. Perhaps the major complaint is not the poor performance but the lack of management control over the manufacturing process.

Many production planning system developers have attempted to improve the quality of their planning and control by storing and processing ever-increasing amounts of data which describe the status and projected future condition of their manufacturing operation. And the bewildering variety of end products and the complexity of the associated sub-assembly

and parts fabrication operations which characterize contemporary manufacturing operations can provide truly enormous amounts of data to store and process. Manufacturing planning and control has been one of the primary users of data processing and, if current trends continue, this will continue to be the case.

Bigger systems processing more data in greater detail have not been the answer. Production planning continues to be a major problem in many firms. Not only is production planning costly, it is not very effective. Inventories are high, parts shortages are endemic, and customer service is not as good as desired. More and more, because of the criticality of production planning to the health of the manufacturing operation and the criticality of the manufacturing operation to the health of the enterprise, not only manufacturing managers but general managers and chief operating officers are turning their attention to production planning systems.

Because of the bewildering complexity and confusion of manufacturing operations, most managers do not want to get involved in the details of production control. They seek general structures which provide a means for delegating successively more detailed decisions to lower levels of management while maintaining the ability to control the overall direction of process.

A division manager wants his plant managers to set the employment rate and overall output rate of the plants for which they are responsible. He also wants the shop superintendents and foremen to deal with specific job and task assignments. However, in doing this, the division manager wants to be confident that the decisions made at these subordinate levels are carrying out corporate and division policies. Clearly, this has eluded most manufacturing managers thus far. As their information systems have become larger and more complex, the details seem always to be further from their grasp instead of being under better control.

There is a way to bring the details under control, i.e., make sure that detailed decisions further further divisional and corporate policies

and objectives, while at the same time delegating the detailed management to foremen and their production control assistants. Several firms have developed and implemented hierarchical production planning systems which are both simple and effective. These systems are based on the notion that senior managers should deal with aggregate decisions based on aggregate data and that shop foremen and their production control assistants should deal with detailed decisions based on detailed data. The top level decisions guide and constrain the decisions made at the next subordinate level. These decisions, in turn, guide and constrain even more detailed decisions until finally the most detailed job assignment and dispatching choices are made.

By separating the decisions to be made into a hierarchy of decisions at several levels, the decisions at each level have a relatively simple structure. One of the keys to the effectiveness of the hierarchical planning approach is the simplicity of the decision making and the associated decision support apparatus.

To illustrate the character of such systems, I will first describe three different approaches to the planning of a reasonably complex production operation. This is followed by a more thorough description of the hierarchical framework, a discussion of the role played by MRP in such systems, and a description of how the recent Japanese concept of "just-in-time" production also fits into the hierarchical structure.

An Example Situation

Let us first look at a real production situation and some alternative approaches to development of the production plans. In this way, we can describe Hierarchical Production Planning in practice before attempting a formal definition.

The firm produces tires, primarily for the replacement market. There are four geographically distributed plants (east, north central, south central, and west). Not all products are produced in each plant. The high volume items are produced in all plants, slow-moving items are produced in only one or two plants. The choice of number and location of

plants for a particular item is a production/transportation economic tradeoff question -- if the volume in a particular region is large enough, it will be cheaper to tool up to produce the item in the region than to ship it in from another plant.

Associated with each plant is a distribution center. It receives products from the local plant and from the other plants (for those items not produced locally) and ships mixed loads of items to customers in vehicle-load and less-than-vehicle load lots.

In the production process, tires are produced in batches in vulcanizing presses. These presses are the most expensive part of the plant and therefore control capacity. Tire companies attempt to keep these presses occupied as much as possible, manning them for three shifts during most of the year. Demand is seasonal and companies do not normally invest in enough presses to allow them to meet the peak demand directly from production.

The overall process is rather complex involving the coating of fabric with rubber to make ply material, coiling of wire into beads, extrusion of rubber into the tread block and assembly of all these components into "green tires." The green tires are cured in the vulcanizing presses. The process in controlled by the curing schedule, the assembly schedule is slaved to the curing schedule, and the coating, bead preparation and extrusion schedules are arranged to satisfy the assembly schedule requirements. Although there are several steps in the process, the basic production scheduling problem is the preparation of the schedule for the curing presses, as though it were a single stage process.

Planning production in this situation is a complex process. Decisions must be made about the production lot sizes of individual items, amounts to be shipped to other plants must be determined, and the location(s) of production for all items must be decided. All these plans and decisions have to be made in the face of a seasonal demand and limited production capacity so that the production planners are usually either building seasonal inventory or using up inventory built earlier.

A major problem is the competition for limited capacity during the peak selling and shipping season. Often it seems that the wrong things were put into the seasonal stock and, as a consequence, the only way the current demands can be met is by using many short runs to make at least small amounts of each of the items in imminent danger of running out.

The Old-Fashioned Approach

Before the day of economic large-scale data processing machinery, these questions would have to be handled on a decentralized plant-by-plant basis. Each plant planner would attempt to build the right seasonal stock of the items produced in that plant, usually concentrating the stock in the high volume items he was confident of selling. This meant that many short runs of slow-moving items were left for the peak season when capacity was at a premium.

Orders were placed by each plant on the other plants for the items which were not produced locally. These orders were placed as late as possible so as to keep the local inventory as low as possible. This led to a high volume of small orders during the peak season, aggravating an already difficult situation.

The essence of this system is that it is decentralized and primarily customer order driven. After an initial centralized plan determines which items will be produced in each plant, each plant manager is pretty much on his own to produce the stock needed to fill the orders he receives from customers and the other regional managers. The primary problem with this approach is the relatively large inventory which is needed, primarily because of difficulties in planning and interplant coordination.

Centralized Monolithic Planning

Watching the "old fashioned" system in operation, it seems clear that large reductions in inventory could be made by improving the "coordination" among the plants. In particular, it seems that small requirements for slow-moving items that are to be produced in one plant

and shipped to another could be produced and shipped during the slack season instead of the peak season. Just how this should be done is not easy to figure out.

The amount of effort required to achieve this coordination is not small. In effect, it requires the development of a detailed production plan for each item to be produced in each of the four plants to satisfy demand anywhere in the system. This plan must cover the seasonal buildup and depletion of inventory and therefore must cover a twelve- to fifteen-month planning horizon. The complete plan will include several hundred thousand variables describing the production of each item each month in each plant, the inventory levels in each distribution center, and the quantities shipped from plants to distribution centers.

There are major problems with this approach. First of all, it is almost impossible for the general manager to review the plan in detail because of the number of variables. Summaries can be prepared but if those indicate that some change is required, it may be impossible to modify the plan in detail so as to yield the desired aggregate result.

If the item plans should add to the total planned production hours during a month, which items should be cut and by how much if the total in some month is one thousand hours too much?

The second problem is the requirement that the planning be centralized with item production plans for each plant being made in a central planning apparatus. This takes away planning authority from the plant managers and reduces their responsibility for plant performance. Finally, the planning apparatus requires a great deal of data and computing power.* It is feasible with available data processing machinery but requires item monthly demand forecasts twelve to fifteen months into the future. These forecasts, of course, cannot be very reliable.

---

* Such a system is described by R.L. King and R.R. Love, Jr., "Coordinating Decisions for Increased Profits," Interfaces, Vol. 10, No., 6, December 1980, pp. 4-19.

Hierarchical Planning Approach

A third approach to the planning process divides the planning process into parts corresponding to the major management levels, using aggregate data for the high-level and intermediate-level decisions and using detailed or item data only for the most detailed scheduling decisions made at the shop level.* The process is illustrated in Exhibit I.

At the corporate level, the only decisions required are those which allocate production among plants. This decides which plants will be equipped to manufacture each item and which distribution regions will be served by each plant producing an item. Month-by-month demand forecasts are not required, annual totals are sufficient.

When production has been allocated to a plant, the seasonal demand pattern has been established. However, the seasonal pattern need not be known in detail for each item; it is sufficient to know the aggregate demand each month for each group of items with a distinctive seasonal demand pattern. The seasonal plan can be prepared for these "aggregate" items and the plan later broken down (disaggregated) into short-term plans for the individual items which make up the aggregate. For instance, one aggregate item was nylon cord snow tires.

In the example case, only five types were required to describe the seasonal demand patterns: nylon cord snow tires, fiberglass belted snow tires, nylon regular highway tires, fiberglass highway tires and light truck tires. Therefore, a very simple seasonal planning system could be used. A linear programming system with a few hundred variables accomplished this. A plant manager, perhaps using a simulation to assist in the process, could also have done this judgmentally.

The seasonal plan is primarily a manpower plan, a description of the number of people to have on the payroll and the planned amount of overtime. It also shows the seasonal inventory build-up and depletion

---

* A.C Hax and H.C. Meal, "Hierarchical Integration of Production Planning and Scheduling," Chap. VI in Logistics (M.A. Geisler, ed.), North Holland, New York, 1975.

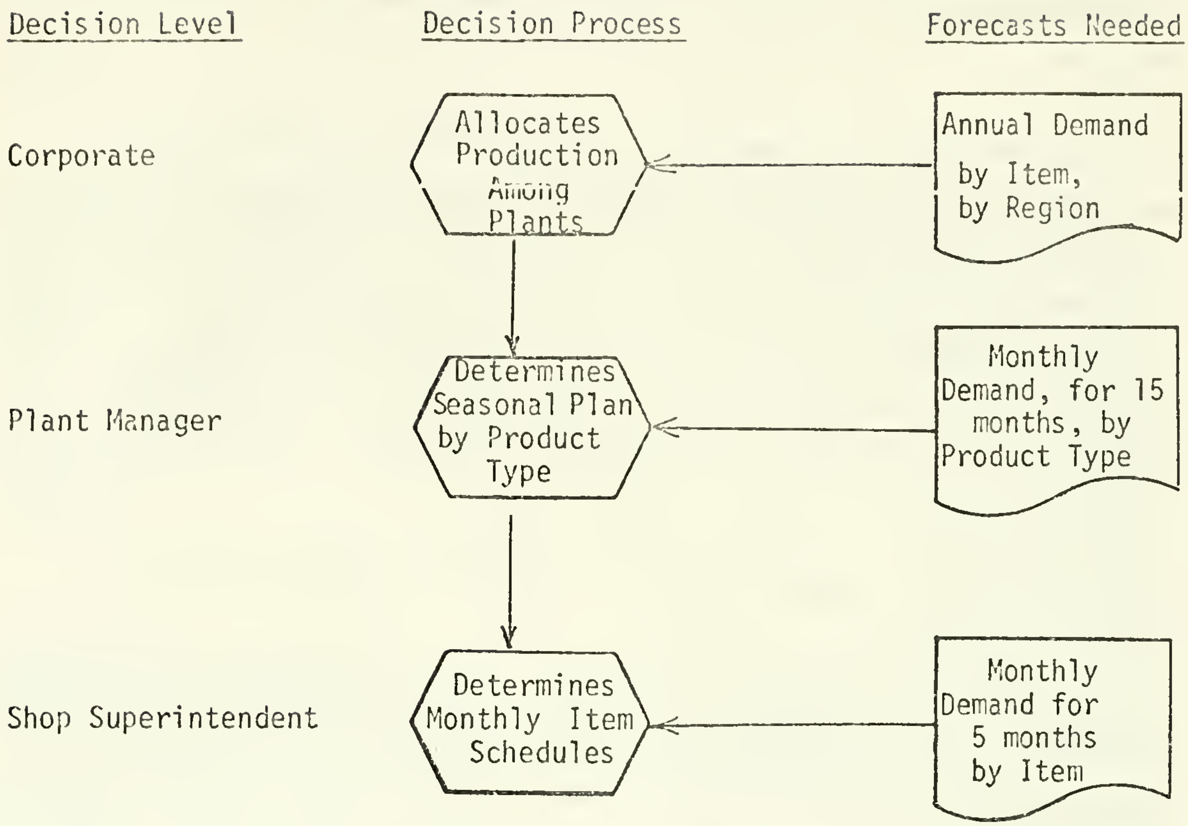

Exhibit I. Hierarchical Planning Process Example

for each of the product or item aggregates. The manpower plan and the resulting aggregate inventory are the primary concerns of the plant management.

Once the manpower plan has been established, it serves as a constraint on the aggregate output of the plant and, therefore, on the schedules of the individual items to be produced. Not only must the total manpower required by those schedules not exceed the total manpower planned, the schedules should use all of the planned manpower. Thus, the final step in this hierarchical scheme is to prepare the detailed schedules in a way which satisfies current customer demand and also satisfies the overall output (manpower) constraint. This is a relatively simple allocation of the available time among the items in each of the types (seasonal patterns) used in preparing the seasonal plan.

For instance, one way to allocate the time available to a particular tire type is to first make sure that all items which are approaching their safety stock levels are included in the schedule. Next, the remaining tire items are produced in quantities sufficient to fill the available production time and also to equalize the time supply of the finished goods inventory for each item.

Using this approach the production planning proceeds in three distinct hierarchical steps:

--allocation of production among plants by corporate management;

--seasonal planning of product groups for each plant by plant management;

--detailed scheduling of individual items at each plant by shop management.

Comments

The hierarchical approach retains the conceptual simplicity of the conventional or old-fashioned approach. Decisions which must be made at the corporate level are centralized; those which can be made locally are delegated to the plants and lower. However, the hierarchical approach

differs from the conventional approach in that the constraints on lower-level decisions which are established by higher-level decisions are explicitly recognized by the decision support apparatus.

As far as the curing room foreman was concerned, the major difference between the hierarchical approach and his former method was in the schedules he received from production control. Under the old system he had a list of items needed; this was developed in the monthly inventory review. This list was updated with a weekly "hot list," items which had run out or were about to. He also had a list of planned total annual production for each item together with cumulative year-to-date production of each.

In the slack season the foreman would never have enough current requirements to use up his available capacity; in the peak season he was always overloaded. In the slack season he used his annual requirements list to select items to fill the available capacity; in the peak season it always seemed that he had selected the wrong things to do in the slack season.

Hierarchical planning made things a lot clearer and simpler. First, it told the foreman how much time to spend on his major tire groups (highway, snow, nylon, fiberglass, light truck). Much more importantly, the system would sort through each of those groups and develop a list of items and quantities to produce which would fill up the available capacity (plant management provided production control with the monthly planned hours) but would also accumulate (or deplete) the seasonal stock in a reasonable way. Both fast and slow moving items were used to accumulate seasonal stock. Some shipments of single plant items to other distribution centers could take place in the slow season, relieving dock congestion during the peak season.

The key to this system is the use of the high level decisions, which allocated production among plans and set the seasonal production plans, to constrain the detailed scheduling decisions made for each item in each plant.

The monolithic approach requires the construction and maintenance of a very large data base and computing procedure. The results are difficult to review since very detailed decisions in each plant are analyzed centrally. In contrast, the hierarchical approach tailors the decision support apparatus to the decisions to be made at each managerial level. Also, since decisions are made in a hierarchical sequence, each decision has a relatively simple structure and does not require large and complex analytical machinery for support. Indeed, the keynote of the hierarchical approach is simplicity.

Hierarchical Production Planning

Hierarchical Production Planning is the division of the entire production planning process, all the decisions to be made, into a hierarchy of decisions to be made sequentially, with the results of the higher-level decisions explicitly constraining the lower-level decisions. The hierarchy is designed to fit the organizational structure and to provide for ease of review at each managerial level. Higher-level decisions have longer lead times, longer planning horizons, and are concerned with aggregates (e.g., total manpower, total product line demand) while the lowest-level decisions have shorter lead times, shorter planning horizons, and are concerned with individual items, machines and workers.

The overall planning process is divided, as nearly as possible, in a way which makes the result of the sequential, hierarchical process nearly as good as could have been obtained from a monolithic process which treated all of the detail in making the most aggregate decisions.

There is a natural hierarchy of decisions ranging from strategic planning choices through tactical planing to detailed or operational planning and control.* This hierarchy is based on the lead times to

*Robert N. Anthony, Planning and Control Systems: A Framework for Analysis, Harvard Univesity Graduate School of Business Administration, Boston, 1965.

execute decisions, the planning horizons to be considered in analyzing and evaluating alternatives, and the magnitudes of the costs affected by the decisions. The hierarchy is natural in the sense that long lead time decisions necessarily constrain short lead-time decisions. For example, the number of people in the workforce is limited by the plant size and the amount of equipment and can be changed (ordinarily) with a shorter lead time than is required to change the plant facilities. Also, the cost differentials resulting from different plant size decisions tend to be much larger than the cost differentials resulting from differences in workforce size.

One of the consequences of these characteristics of the hierarchy of production planning decisions is relative independence of decisions. That is, high-level decisions can usually be made without taking into account the way in which lower-level decisions will later be made. One can use aggregate instead of detailed data in making aggregate decisions and incur little, if any, cost penalty as a consequence of the approximation.

For example, one can decide on the manpower level without deciding the detailed lot sizes for the items produced by the manpower during a month. Of course, we need to know enough about the way the detailed schedules are prepared to know roughly how much time is lost to changeovers, so that can be planned for. That is, we need to know the aggregate behavior of the process when making aggregate decisions about that process. The aggregate decision, however, does not depend on the detailed decisions that set lot sizes. In fact, it is exactly the opposite. The detailed lot size decisions will be made later, when better requirements information is available and will be constrained by the aggregate manpower (output) decision. The lot sizes will be selected so as to use exactly the manpower available and satisfy requirements as well as possible.

A very general kind of hierarchy is shown in Exhibit II. This is an extension of Exhibit I showing the plant sizing and location decisions and the equipment selection decisions. These decisions all are required

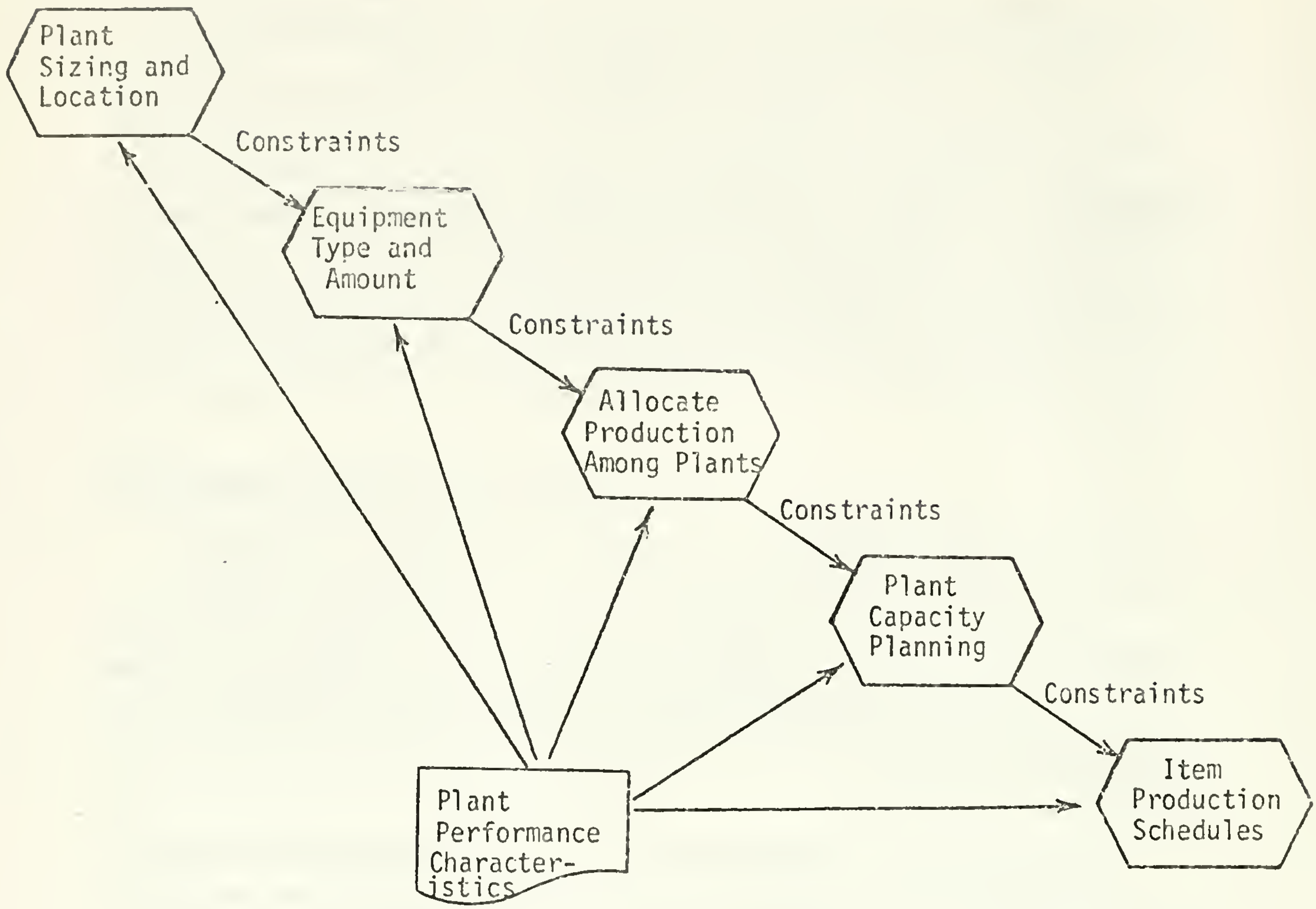

Exhibit II. General Production Planning Decision Hierarchy

to complete the preparation of a production plan. Clearly, the decisions at the top constrain those below in a sequential fashion. They have longer lead times, longer planning horizons, make use of more aggregate data, and involve larger amounts of resources. Also, they characteristically are made by higher levels of management. Major plant and equipment investment decisions are made (or at least reviewed by) the Board of Directors. Allocation of production among plants is done at the level of the Chief Operating Officer, seasonal plans are prepared by plant manager, and item schedules are prepared by the shop superintendents and foremen. These very detailed decisions are highly constrained by all that has gone before and should be guided by corporate policies. The shop production planners need to be provided with the appropriate tools to help them implement the corporate policies within the constraints that exist.

Each level of decision making obtains feedback of aggregate performance information from the process itself. That information reflects the effects of detailed decision making.

Illustrative Example

A major manufacturer of consumer non-durable goods faced many problems which were similar to those of the tire manufacturer described earlier. Demand was seasonal and the costs of building seasonal inventory prevented use of a completely level work load plan. However, the company had a tradition of employment stability and was reluctant to hire and fire in response to the seasonal changes in demand.

There were no short range problems in allocation of production among plants since nearly all products were made in only one plant. The few very high volume products that were made in two plants were allocated geographically so as to minimize transportation costs and could be treated as single plant products.

The problems were very similar to those of the tire manufacturer. The most striking problem is one of dealing with the large number of items produced by a particular assembly line. The items compete for

production time and there were problems in trying to meet peak demands just as there were problems in trying to figure out what to do during the slack season.

Production planning was clarified and simplified by putting together a planning process in which the general manager presided over the development of the aggregate output plan. This plan addressed directly the tradeoff between total inventory buildup and depletion and the manpower plan at each assembly plant.

Once this plan was prepared (quarterly) the plant managers knew the constraints within which they would operate. Production control prepared schedules for the items which were constrained to fill, but not exceed, the available time. These schedules were perhaps the primary benefit to the shop foremen and the plant managers. They were feasible and economic, and they yielded a balanced inventory of finished goods.

In the planning process the general manager could be confident that his inventory projections (in the aggregate) would hold if the aggregate demand forecast was accurate. Also he had a good idea of the error distribution. The general manager was also comfortable that his plant managers knew what they had to do to make the company's aggregate results match the budget. Finally, this framework of guides and constraints greatly simplified the life of the foremen and the production control assistant who could plan line changes and order parts with reasonable confidence they would not be called upon to interrupt runs to deal with "emergencies."

Structuring the Hierarchy

The correct hierarchical structure for a given firm depends on the character of the products, the production technology, the costs associated with various actions, and the organizational structure of the firm. For the most part, when we are discussing production planning, the plant and equipment decisions have been made and we are concerned with structuring the remaining routine or repetitive decisions to be made about production plans within the context of fixed plant facilities.

In structuring the hierarchy, there are often choices to be made about the level of detail to be included in a particular decision, for example, should the seasonality of demand (if any) at the several plants be taken into account when deciding how to allocate production among plants? The general rule to follow is to include as little detail as possible without incurring significant error. If the plant allocation can be made without much error while ignoring demand seasonality at the plants, by all means do it that way. The simpler the decisions, the better. In some cases, demand seasonality is important in plant allocation, in other cases, not.

From this we see that the hierarchical structure may be, and often is, highly idiosyncratic. The structure which is best for one firm in an industry may not be the best for another firm in the same industry. There are some guidelines, however, to follow in structuring the hierarchy.

Keep the decisions simple. Use aggregate data whenever possible and limit the number of variables which have to be considered simultaneously.

Separate decisions if possible. Look for decisions which have little, if any, effect on each other and make sure they are made separately. Usually lot-size decisions made about a particular item in one plant have no impact on the correct lot-size choice for the same item in another plant. Make sure the lot sizing rules recognize this.

Match the Decisions and the Organization. The organizational structure in most firms is reasonably sensible. Since we need to identify who will be responsible for each of the decisions in the hierarchy, it is often helpful to identify which decision is to be made by each level in the existing organizational hierarchy. Often this will help show how the planning hierarchy should be structured. Occasionally it will become clear that some change is needed in the assignment of responsibilities. This happens when it becomes clear that either the needed information or the required authority is not available to the person currently charged with making a particular decision.

Even with these guidelines, it takes a substantial amount of observation and analysis of the operation and the existing production planning process to arrive at the right structure for a hierarchical production planning system. When the right hierarchy has been found, it should be fairly clear that it is the appropriate structure. As Bob Brown* once said, "The right answer may not be obvious, but it should be obvious that it is right."

Planning of Multi-Stage Operations

The strength of the hierarchical approach shows even more clearly in multiple-stage operations, particularly in those with multiple, specialized parts plants and assembly plants. These are characteristic of the automobile and computer industries, to cite two common and very diffierent examples.

The hierarchical approach works well in multiple stage production because it facilitates the interstage coordination. This may be seen in Exhibit III, which represents the material flow and storage through successive stages of parts fabrication, assembly, and distribution, and also shows the essentials of the planning processes. The parts fabrication block should be understood to represent several or many parts fabrication units or suppliers. Similarly, the assembly block represents several assembly units, operating in parallel and producing a wide variety of finished products. In one large organization where this approach has been implemented, there are more than fifty assembly units in seventeen different assembly plants and similar numbers of parts fabrication units and plants. Parts move from each parts unit to several assembly units and each assembly unit draws parts from several parts fabrication units.

Only two levels of the planning and control hierarchy are shown in Exhibit III. These are the planning and control of the aggregate output of each unit and the detailed planning (item lot sizes and timing) of

---

* Robert G. Brown, a former colleague, is now President of Material Management Systems, Inc., Norwich, VT.

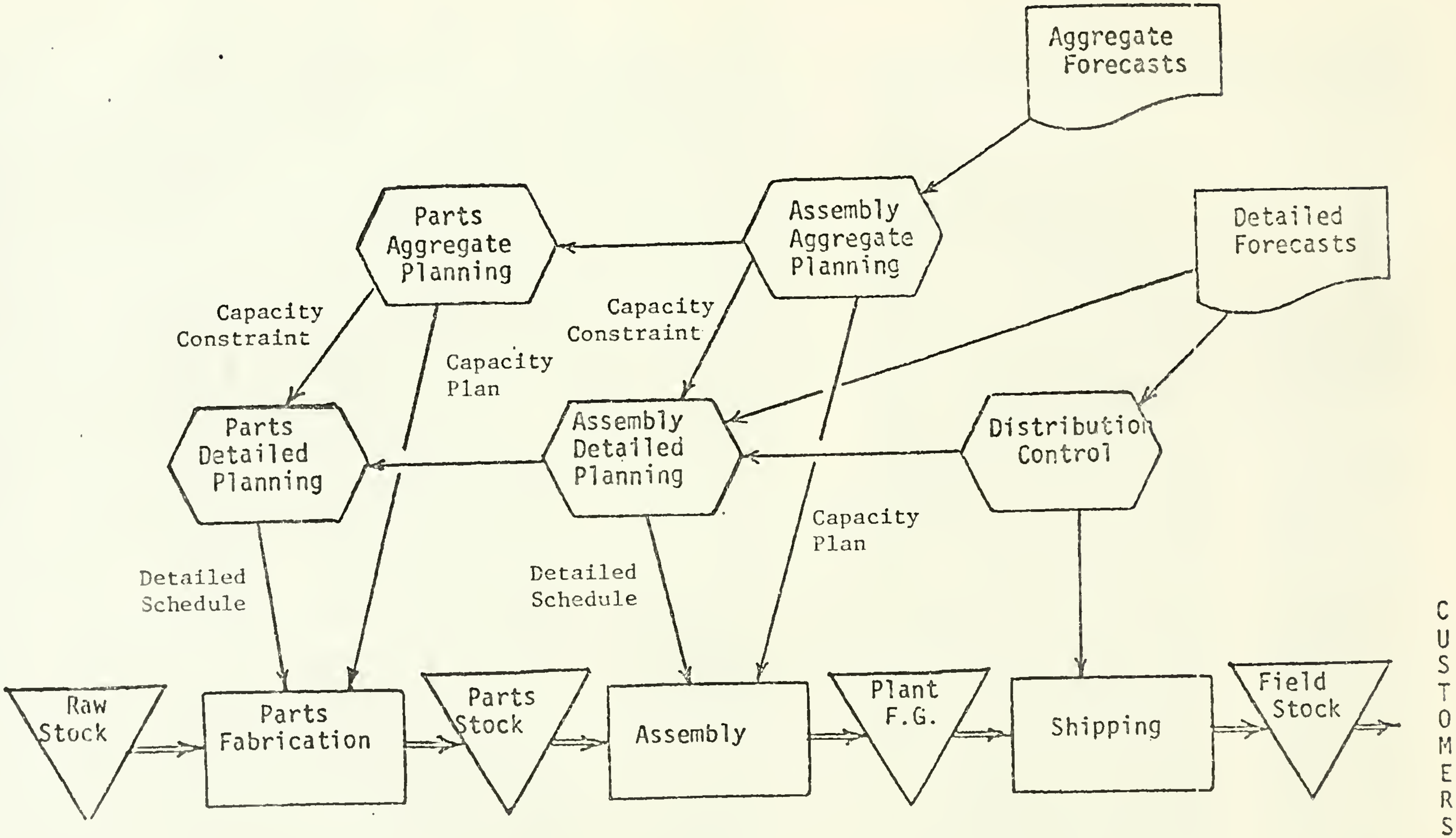

Exhibit III. Multi-Stage Hierarchical Planning System Structure

each parts and assembly unit. This assumes that plant sizing, locating, and equipping decisions have already been made and constrain the choices here. Any allocation of production among plants with similar or interchangeable capability has also already been done.

At the aggregate level, it is important to have all the stages in the operation running at the same rate. If they are not, excess parts inventories will build or, conversely, the assembly operation will be starved for parts. Similarly, lack of correspondence in operating rates of the assembly and distribution activities will lead to finished goods inventory problems. These problems are aggravated when there is a general increase or decrease in demand, characteristic of the short life cycles of high technology products such as computers. In those cases, the early stages in the operation must anticipate the later stages, building up and cutting back ahead of the final demand.

At the aggregate level of control, operating rates are set. A central planning activity can set these rates for the entire system. Alternatively, the assembly rate can be set (on the basis of a demand forecast) and all prior stage rates derived from the assembly rate. This determination can be based on conversion factors relating the manhour (or other critical resource) requirements in the early stages to the aggregate output rate of the assembly units. This has a strong similarity to Material Requirements Planning (of which more later), but differs in one crucial respect. In MRP, one works with a specific assembly schedule for finished products and uses the assembly bill of materials to develop prior stage <u>detailed</u> requirements. Here we are dealing with <u>aggregate</u> output rates and therefore do not have bills of material to translate the aggregate plan at one stage into requirements at an earlier stage. The notion of a "planning bill of materials", a listing of the aggregate resources required to complete each aggregate unit of output is used to assist in aggregate planning. The bill can be as simple as a conversion factor which describes the manhours required in a subassembly unit for each thousand dollars in sales of a particular product line.

The great strength here of the hierarchical approach lies in the availability of an explicit statement of aggregate output which can be used

to guide the output planning of an earlier or later stage. Normally, these plans are made one to three quarters in advance with the longer lead times being common if a substantial amount of training is required to develop the required operator skill level.

The detailed schedules for both the assembly and parts fabrication processes are made with much shorter lead times—two to four weeks being common. Often the master assembly schedule is known far enough into the future that the requirements for parts during the parts production lead time can be determined by exploding the master assembly schedule into its component parts. This is the technique called Material Requirements Planning (MRP) and which provides an excellent forecast of requirements when the assembly schedule is known far enough into the future.

Even if MRP is not used to develop the earlier stage requirements, the wisdom of planning with two levels should be apparent. The aggregate level provides consistency in running rates for all stages in the process; the detailed level makes use of the most recent demand information and inventory status information in allocating the available production time among the individual items to be produced.

The consumer durables manufacturer described earlier manufactured some of his own parts in separate plants. Each of these parts plants would serve more than one assembly unit. Prior to taking the hierarchical approach these parts plants were controlled more or less independently; they had dollar budget figures from the assembly plants but no reliable operational planning data. Instead, the assembly plants placed orders on the parts plants, more or less as though they were outside vendors. The parts plants responded as best they could and had the usual amount of trouble because of order changes, over-ordering, under-ordering, second guessing on both sides, etc.

When the hierarchical system was installed the parts plant manager was able to set his operating rate in response to the planned operating rates, by product group, in final assembly. Some months he ran a little ahead, some a little behind, but he had a much more orderly and economic process than when he tried to set his operating rate on the basis of orders received.

Now, when orders come in, he can normally fill them from the production planned for the period or from inventory. Orders in hand and the inventory status are used to assist in preparing the item production schedules within the overall (aggregate) operating rate constraint.

MRP in Hierarchical Production Planning

MRP is a technique for developing, at any stage in the production process, the requirements which are a consequence of the final stage detailed master schedule. MRP can equally well be described as a simulation of the parts fabrication and sub-assembly activities that are required, in the appropriate time sequence, to accomplish the assembly master schedule. If the master assembly schedule is known for a time in the future long enough to accomplish all the parts procurement, parts fabrication, and sub-assembly activities, a deterministic requirement statement is available. If, in addition, capacity at all the prior stages can easily be adjusted to match the manufacturing requirement, all these production activities can be carried out with very little requirement for buffer inventories between stages.

Fluctuations in the requirements from one period to the next present a problem in capacity management. Often, the period to period variation is substantial. Even when the total load in final assembly is reasonably level, variations in the product mix will lead to variations in load at particular manufacturing centers. Run length economics may require combination of two or more periods' requirements into a lot to be produced in one period, increasing the period-to-period fluctuations still more.

The total load on a particular subassembly or parts manufacturing center is the sum of the requirements of the several or many individual items produced by the center. Some averaging takes place, but the normal state of affairs is large period-to-period variations in total load on the center. Every shop foremen is familiar with this "feast or famine" situation, and does what he can to deal with it. Economy of production dictates that these fluctuations be smoothed to some extent. MRP provides no guide for this smoothing process, it just states the requirement.

Production smoothing is the result of the aggregate output planning (capacity planning) process. In most production systems this is the next higher level in the planning structure. Thus, in the hierarchical planning approach, smoothing the requirements or load at each manufacturing center is a natural consequence of carrying out the planning process. When the capacity plan for each center is established, detailed planning (production scheduling and dispatching) can proceed. The objective of the detailed planning is to meet the requirements statement economically and within the constraint of the (smoothed) capacity plan for the center.

Hierarchical planning does not guarantee that it will always be feasible to meet all the requirements, although providing some smoothing inventory can keep the number of problems at a manageable level. It does provide an explicit comparison of capacity with requirements (as does the capacity requirements planning module in advanced MRP systems) and also provides a scheduling method which makes best use of the available capacity.

MRP does not address the question of how the schedule at each stage should be prepared. Most commercial MRP systems assume there are no capacity limits in deciding on the schedule. These systems use unconstrained lot sizing methods to prepare the schedule.

If the capacity (usually labor hours) of the subassembly and parts fabrication centers can be changed quickly and economically, to match the swings in requirements, use of the unconstrained lot sizes is appropriate. The capacity is matched to the load and the requirements are met. However, tracking the fluctuations in load with variations in work force or overtime may not be economic. It may not even be physically possible.

In operations involving a high degree of mechanization or automation, or requiring highly trained technicians, it may not be possible to accommodate wide load swings. The scheduling methods must reflect this characteristic of the operation. Constrained scheduling methods must be used. The MRP system can readily use such methods at any center requiring them. When that is done, we really have the best of both worlds. MRP provides the most

accurate statement of the requirements to be satisfied; the hierarchical planning approach provides the framework in which to use the appropriate scheduling method.

In the "closed loop" version of MRP, the capacity requirements at each stage are worked out and compared with the available capacity. If there is sufficient capacity, no problem. If not, the master schedule is modified in an attempt to arrive at a feasible schedule for all stages. This modification must be done intuitively since we do not yet have analytical methods which will tell us how to reduce the master schedule by the amount required to just make feasible the schedule of an earlier operation. The version called MRP II incorporates the capacity checking feature as well as other extensions of the basic concept.

Some companies attempt to accomplish a "bottom-up" capacity planning technique. They extend the master schedule out far enough to provide a capacity requirements estimate for the several quarters needed in planning capacity of parts fabrication operations. Since the later periods in the master schedule are quite likely to change, there is substantial uncertainty in this requirement statement, but that is not the most serious problem. The primary problem here is that of summing the detailed requirements to provide the aggregate estimate. Because it is based on all this detail, it has an apparent validity which exceeds its actual accuracy. Nearly always, a better capacity plan can be obtained by starting at the aggregate level and developing the plan directly.

In contrast, hierarchical planning avoids these problems by using a "top-down" approach.

Just-in-Time Production

Among the practices "discovered" in recent Japanese manufacturing practices is that of reducing the in-process inventory to as low a level as possible by reducing the lot sizes and arranging a communications and control system which brings the replenishment lot to the place it is needed "just in time" to avoid a disruptive run out. One of the elements in this system is the "kanban", meaning "card", or "tag", or "shop symbol."

Whenever a container of parts is used, a "kanban", attached to the container and describing the part, is detached and returned to the supplier as a signal that a certain amount of material has been used and, therefore, an additional quantity should be provided. The supplier receives the kanban and attaches it to a full container of parts which he dispatches to the manufacturer. The total number of kanbans in the system is equal to the number of containers of parts in transit from the supplier, plus the number of containers (usually just one) waiting at the assembly plant and in the process of being emptied, plus the number of kanbans enroute back to the supplier. If the usage rate is one container per day, the transit time from the supplier is five days, and the time to return the kanbans is one day, there will be seven kanbans circulating and controlling the flow.

By controlling the number of kanbans in the system, the timing of replenishment container arrivals can be controlled. Adding a kanban will bring the arrivals earlier; taking one kanban out will make the arivals later. By reducing the size of the containers the inventory on hand at the assembly plant can be reduced. Of course, the number of kanbans must be increased to compensate.

This is a simple, effective system for controlling the levels of in-process inventory in a system operating at a constant or slowly changing overall rate. It deals with short-term fluctuations in usage, such as those characateristic of option application in automobile assembly. (It is not new, having been described in the early 1950's in this country. It was then called the "base stock" system.) This system works well when the overall rate is constant but is unsatisfactory for communicating basic changes in production rate to the earlier stages in the process. The time delay in the system as a consequence of the time required to accomplish the earlier stages in production require that these earlier stages start producing at the new rate <u>before</u> the final stage changes to that new rate. In the kanban or base stock systems, the notice of the higher production rate is sent back to the early stages only at the time the final stage changes to the new rate. If the new rate is higher, the available material will quickly be exhausted.

On the other hand, using the hierarchical approach, the earlier stages do not rely on these short-term signals to establish their production rates. Instead, with a relatively long planning horizon, these early stages look at the planned future final assembly rates to tell them what they need to use as their operating levels. This includes any anticipation of increases or decreases in the assembly rate. With the parts fabrication and subassembly rates set in this way, the kanban system serves as a very effective method to move parts to the subassembly areas and subassemblies to the final assembly lines. It is very simple to add (or subtract) kanbans to the system to increase (or decrease) the pipeline inventory.

This situation is very similar to that of MRP. The aggregate output planning provided by the hierarchical production planning structure is the framework for the detailed planning and control accomplished in part by MRP or just-in-time supply of parts. Without that aggregate planning to provide a context for the detailed planning, neither of these two detailed approaches is very effective.

Manufacturing Control and Performance Measurement

As manufacturing operations have become more complex and disperse, directing the operation in a coordinated way has become more difficult. With conventional planning methods it is difficult to tell the managers of the upstream plants and manufacturing units how they should operate to support a particular product mix for the company as a whole. As the final output mix changes there is no convenient way, using conventional methods, to guide the several manufacturing unit managers in making their manpower, facilities, and procurement plans.

Later, in the execution of those plans, the lack of coordination shows up at the detailed, or item, level. We have too many of one part or group of parts, too few of another. One plant may be operating at too low a rate and another at too high a rate. Since their capabilities are different the excess capacity cannot fill in the capacity shortage. Part shortages (and excesses), delays, and expediting are the result.

One of the common reasons for this class of problems is the practice of preparing operating plans in financial terms. Even if finished product

plans are separated by product line, there is usually no way to relate the product line mix changes to changes in the dollar mix at the several supporting manufacturing units.

The coordination and delegation needed in a complex manufacturing control system is relatively easy using hierarchical production planning. Since the system is top-down it is natural to coordinate plans at the top level before breaking them down into detailed plans. This ensures that all parts of the operation are operating at the desired rate before getting involved in details that are very tedious to try to coordinate.

Delegation is much clearer with HPP. Each unit manager knows what his operating rate is supposed to be, and, at successively more detailed levels, he knows where to obtain the requirements information needed to assist in allocating this operating rate among production units, machines or items.

The operational planning system parallels the financial planning system. In conventional planning structures operational planning and financial planning are often in conflict. Sometimes this happens because the detailed plans do not add up to the desired aggregate plan. With HPP this problem is avoided; the financial plan and the operational plan have the same structure and tend to stay in correspondence as they are both disaggregated.

In measuring operating unit or unit managerial performance, that performance can be compared to the operating plan rather than the financial plan. This reduces the incentive to do things that improve financial performance of the unit but reduce the operational performance of the company. This follows directly from the fact that HPP clarifies and simplifies level to level delegation and stage to stage coordination.

Implementation of HPP Systems

One of the great strengths of the hierarchical approach is its ease of implementation. Since it deals with reasonably separable decision problems in a sequential fashion, it is natural to implement the new decision support

systems sequentially instead of all at once. This is much easier and less risky than attempting to implement a system which deals with all decisions in a single monolithic process.

The first step is the design of the hierarchical structure. That design should include consideration of implementation questions. Ideally, the several levels in the structure will be nearly independent so that one level can be modified (implemented) with a minimum effect on the other levels. A natural way to sequence the implementation is to go from the top of the hierarchy to the bottom. Often, this will implement improved decision making in a sequence which provides the greatest benefit earliest in the program. Conceptually, this is a satisfying sequence to follow since each time a decision support system is to be implemented, we find already in place the system which provides the needed constraints.

Conclusion

Production planning systems need not become more and more complex as product lines proliferate and manufacturing processes are subdivided into multiple stages. Structuring the production planning process in a hierarchical way allows the use of simple decision support systems using aggregate data for high-level decisions and detailed data for shop-floor decisions. Tailoring the planning structure to the organization along with keeping the decisions simple permits clarity in delegation and maintenance of overall guidance and control of the manufacturing process by top management.

Acknowledgement

I would like to acknowledge the major contribution to this work of Professor Arnoldo C. Hax, my colleague here at MIT and collaborator in the development of the basic ideas of hierarchical production planning. H.C.M.

Date Due

Bar Code
On last
Page